

Jani Koistinen

**AVOIMEN DATAN SAATAVUUS JA  
HYÖDYNTÄMINEN  
SÄHKÖVERKKOTOIMINNAN  
KEHITTÄMISESSÄ**

Informaatioteknologian ja viestinnän tiedekunta  
Kandidaatintyö  
Joulukuu 2019

# TIIVISTELMÄ

Jani Koistinen: Avoimen datan saatavuus ja hyödyntäminen sähköverkkotoiminnan kehittämisessä

Kandidaatintyö

Tampereen yliopisto

Tieto- ja sähkötekniikan TkK-tutkinto-ohjelma

Joulukuu 2019

---

Tämä Työ tehtiin kirjallisuusselvityksenä ja siinä tutkittiin minkälaista avoimen datan saatavuutta ja kuinka sitä voitaisiin hyödyntää sähköverkkotoiminnassa sekä minkälaista tutkimusta avoimen datan hyödyntämisestä on tehty.

Maailmassa on datan määrän on lisääntynyt valtavasti viime vuosina. Samalla avoimen datan saatavuus Euroopassa on helpottunut huomattavasti kun valtiot ovat avanneet tietokantojaan INSPIRE direktiivin myötä ja ihmisten vaatiessa avoimempaa yhteiskuntaa tiedon suhteen. Tämä avoin data on saanut aikaan uudenlaista liiketoimintaa sen ympärille ja se on ollut myös tiedon avaamisen tarkoituksena.

Avoim data sisältää usein jonkin viittauksen paikkatietoon. Avoin data itsessään ei välttämättä ole arvokasta käyttäjälle. Vaan se usein täytyy muokata ja yhdistellä käyttäjän omiin tarpeisiin sopivaksi. Avoimen datan hyödyntäminen täysimääräisesti vaati myös käyttäjältä teknistä osaamista johtuen juuri sen muokkaamisen ja yhdistämisen tarpeesta.

Sähköverkkotoiminnassa kulutuksen ennustamisessa käytetään vanhoja, pitkään käytössä olleita malleja. Avointa dataa ja nykyisten uudenaikaisten etäluettavien mittareiden tietoja hyödyntämällä näitä malleja voidaan päivittää. Samalla saadan tarkempi kuva miten sähkönkulutus jollakin alueella voi kehittyä ja osataan tehdä valistuneempia päätöksiä investointien suhteen. Avoimen datan voidaan katsoa olevan päätöksenteon apuväline sähköverkkoyhtiöille.

Avainsanat: avoin data, sähköverkko, sähköverkkotoiminta, kehittäminen, paikkatieto

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

# SISÄLLYSLUETTELO

1	Johdanto . . . . .	1
2	Avoim data tiedon lähteenä . . . . .	2
2.1	Avoim data . . . . .	3
2.2	Avoimen datan ominaisuudet . . . . .	4
2.2.1	Löydettävyys . . . . .	4
2.2.2	Kokonaisuus . . . . .	5
2.2.3	Käyttöehtojen tasa-arvoisuus . . . . .	5
2.2.4	Alkuperäisyys ja ajantasaisuus . . . . .	5
2.2.5	Laillinen ja vapaa uudelleenkäytettävyys . . . . .	6
2.2.6	Maksuttomuus . . . . .	6
2.2.7	Koneluettavuus . . . . .	6
2.2.8	Formaatin avoimuus . . . . .	6
2.2.9	Ymmärrettävyys . . . . .	7
2.3	Avoimen datan julkaisu . . . . .	7
2.4	Avoim data Suomessa . . . . .	7
2.5	Avoimen datan ongelmat . . . . .	8
2.6	Avoimen datan tulevaisuus . . . . .	8
2.7	Esimerkki avoimen datan saatavuudesta ja sen esittämisestä . . . . .	9
3	Sähköverkko . . . . .	11
3.1	Kantaverkko . . . . .	11
3.2	Jakeluverkko . . . . .	12
4	Sähköverkkotoiminta . . . . .	14
4.1	Toimitusvarmuusvaatimukset . . . . .	14
4.2	Sähköverkon kehittäminen . . . . .	15
4.3	Smart grids . . . . .	15
4.4	Sähköverkkotoiminnassa käytettävät tietojärjestelmät . . . . .	15
5	Avoimen datan hyödyntäminen sähköverkkotoiminnassa . . . . .	17
6	Yhteenveto . . . . .	20
	Lähteet . . . . .	21

## LYHENTEET JA MERKINNÄT

5G	Viidennen sukupolven mobiilidatayhteys.
AMR	Etäluettava mittari (engl. Automatic meter reading).
API	Ohjelmointirajapinta (engl. Application programming interface).
CSV	Tiedostomuoto yksinkertaisen taulukkomuotoisen tiedon tallentamiseen (engl. Comma Separated Value)
DIKW	Tulee englannin sanoista data, information, knowledge, wisdom.
ENTSO-E	Eurooppalaisten kantaverkkoyhtiöiden yhteistyöjärjestö.
GIS	Paikkatietojärjestelmä (engl. Geographic Information Systems).
HTML	Kieli hypertekstin esittämiseen (engl. Hypertext Markup Language)
INSPIRE	EU-direktiivi paikkatietoinfrastruktuuriin.
IoT	Esineiden internet (engl. Internet of Things).
JSON	Java-pohjainen tiedostomuoto tiedon välitykseen (engl. JavaScript Object Notation)
KNN	K:n lähimmän naapurin menetelmä (engl. K-nearest neighbours).
PAAVO	Postinumeroalueittainen avoin tieto.
PDF	Dokumenttiformaatti (engl. Portable Document Format)
PxWeb	Verkkopohjainen taulukkojärjestelmä tilastontuottajille.
RDF	W3C:n standardimalli tiedon vaihtoon sovellusten välillä (engl. Resource Description Framework)
SOM	Itseorganisoituva kartta -algoritmi (engl. Self-organizing map).
WFS	Rajapinta paikkatiedon esittämiseen (engl. Web Feature Service).
WMS	Rajapinta paikkatiedon esittämiseen karttakuvina (engl. Web Map Service).
XML	Tiedon kuvaamiseen käytettävä kieli (engl. Extensible Markup Language)

# 1 JOHDANTO

Sähköverkkotoiminta on tällä hetkellä murrosvaiheessa, kun säävarmojen verkkojen rakentamisen lisäksi kuluttajat ovat valveutuneet vaatimaan sähköltä myös ympäristöystävällisyyttä. Samoin lisääntynyt sähkön pientuotanto, sekä kasvanut sähköhybridi- ja täys-sähköajoneuvojen myynti vaikuttavat siihen, kuinka sähköverkkoa ja sähköverkkotoimintaa tulisi kehittää.

Samaan sähköverkkojen murrosvaiheeseen liittyy myös älykkäät sähköverkot ja saatavilla oleva data sähkön käytöstä. Työssä tutkitaan, olisiko sähköverkkoyhtiöillä mahdollista hyödyntää saatavilla olevia datavarantoja sähköverkkotoiminnassa. Datavarantojen kohdalla keskitytään nimenomaan avoimeen dataan, sen saatavuuteen ja hyödynnettävyyteen. Avoimen datan hyödyntäminen olisi sähköverkkoyhtiöille kustannustehokkainta, sillä avoimen datan tulisi olla kaikille käyttäjille maksutonta. Avoimen datan monet ominaisuudet käydään läpi luvussa 2.

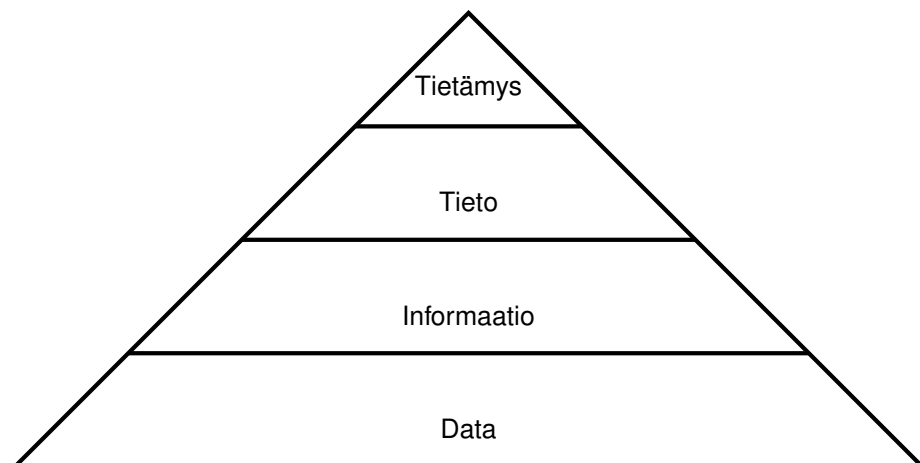
Työssä käydään luvussa 2 aluksi läpi avoimen datan syntyä ja minkälaisia käsitteitä ja ominaisuuksia avoimeen dataan liittyy. Tässä luvussa on myös oma visuaalinen esimerkki avoimen datan hyödyntämisestä QGIS-ohjelmalla. Tämän jälkeen luvussa 3 tehdään lyhyt katsaus sähköverkkoon ja luvussa 4 sähköverkkotoimintaan Suomessa. Luvussa 5 pyritään tuomaan esiin minkälaisia hyödyntämismahdollisuuksia avoimella datalla voisi olla sähköverkkotoiminnassa. Lopuksi viimeisessä luvussa tehdään vielä yhteenveto työssä käsitellyistä asioista.

## 2 AVOIN DATA TIEDON LÄHTEENÄ

Datan määrä on kasvanut viime vuosina hurjaa vauhtia, ja sen koon määrittäminen tarkasti on mahdotonta. Kun arvioidaan datan määrän kasvua vuodessa, puhutaan zettatavun suuruusluokista [1]. Valtavasta datamäärän kuvaamiseen on yleisesti vakiintunut termi Big Data.

Datan tuottamisesta voidaan ajatella, että se on tutkimus ja kehitystyötä, koska tästä saatava rahallinen hyöty ei välttämättä konkretisoidu heti. Datan tuottamisen kustannukset ovat suuret ja siitä saatava rahallinen hyöty on vasta kaukana tulevaisuudessa. Tästä johtuu, että yritykset eivät sitä tahdo välttämättä kerätä, vaikka tarjolla olisi paljon dataa, jolla olisi mahdollista markkina-arvoa. Ilman julkishallinnon panostusta yliopistojen ja tutkimuslaitosten perustutkimukseen, olisi datan tutkimus ja kehitys vähäistä verrattuna nykyiseen. [1]

Puhekielessä sanat data ja tieto hyvin usein sekoitetaan keskenään, vaikka varsinaisesti kyseessä on kaksi eri asiaa. Datan suhde tietoon ja datan arvoketju voidaan hyvin ymmärtää käyttämällä pyramidimallia datan, informaation, tiedon ja tietämyksen suhteista ja järjestyksestä. Käsitteiden suhteet eivät ole pyramidissa vakioita, vaan voivat vaihdella eri kerrosten välillä. [1]



**Kuva 2.1.** DIKW-pyramidi [mukaillen 1].

Kun puhutaan datasta käsitteenä, tarkoitetaan sillä digitaalisesti tallennettua merkkien ja symbolien joukkoa [2]. Datakerroksessa esimerkiksi jokin laite havainnoi ympäristöään tai toimintoa, tallentaa sen ja välittää sen edelleen informaatioksi. Informaatiosta sen käsit-

telijä voi edelleen jalostaa tätä tiedoksi ja tietämykseksi. Data ei itsessään ole välttämättä arvokasta. Datan arvo määräytyykin vasta sitten, kun se kyetään jalostamaan ja hyödyntämään omassa toiminnassa. Datan voidaan siis ajatella olevan kuin jokin teollisuuden raaka-aine, jota tarvitsee muokata ja jalostaa, jotta siitä saadaan irti se tärkein eli arvon tuottaminen

Data määrän hurja kasvu asettaa myös haasteita sitä tuottaville yrityksille. Gartner varoitti vuoden 2014 artikkelissaan [3], kuinka suuret yritykset ovat ongelmissa tietoturvan kanssa. Hyvänä esimerkkinä dataan liittyvistä tietomurroista mainittakoon suurten kryptovaluuttapörssien tietomurrot, kryptovaluuttahuuman ollessa kuumimmillaan. Tämä jo itsessään antaa viitteitä siitä, mikäli datalla on arvoa yhdelle osapuolelle, niin sillä todennäköisesti on arvoa myös toiselle osapuolelle.

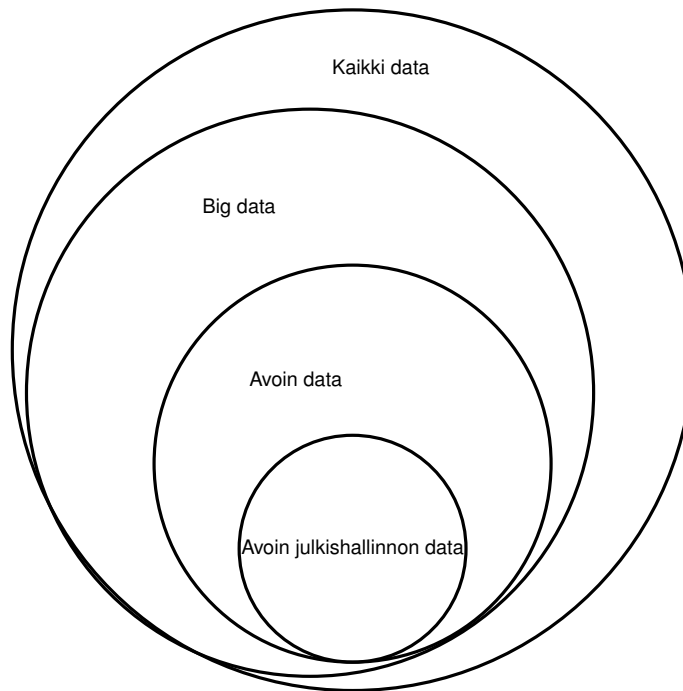
## 2.1 Avoin data

Kauan ennen kuin avoimesta datasta tuli julkishallinnon apuväline avoimuuteen, ajatus tiedon avoimuudesta oli jo juurtunut tiedeyhteisöön tutkijoiden välityksellä, sillä tutkijat olivat ensimmäisiä, jotka jakoivat tietoa keskenään. Idea tiedon avoimuudesta alkoi vuonna 1942, jolloin sosiologi Robert K. Merton esitti, että tieteellisten tutkimustulosten tulisi olla kaikille avoimia, ja että kaikkien tutkijoiden tulisi olla mukana tiedon eteenpäin viemisessä. Avoin data terminä esiintyi kuitenkin ensimmäisen kerran vasta vuonna 1995 Yhdysvaltalaisen National Research Councilin julkaisussa *On the Full and Open Exchange of Scientific Data*. [4]

Tiedon avoimuus on tällä hetkellä yhä useamman ja useamman valtion aiheistalla. Tietojen avaamisen tarkoituksena on ollut talouden kehittäminen, innovointi ja julkisen hallinnon toiminnan tehostaminen [5]. Manyika et al. [5] toteavat raportissaan, että vuonna 2013 yli 40:llä valtiolla oli jo tuolloin jonkinlainen avoimen datan alusta, ja että avoimen datan yhteenlaskettu arvo olisi noin 3 biljoonaa Yhdysvaltain dollaria.

Avoin data ajatellaan usein olevan jokin julkishallinnon alainen ilmainen palvelu. Avoin data voi myös olla yhteisön tai yrityksen tarjoamaa ja se voi olla tietyn ehdoin hyödynnettävissä olevaa tai kokonaan maksullista. Avoin data on mahdollistanut sen, että dataa ei tarvitse enää itse välttämättä tuottaa ja kerätä vaan siitä vapautuneet resurssit voidaan käyttää kokonaan sen hyödyntämiseen. Avointa dataa hyödyntämällä on mahdollista tuottaa uusia palveluita, tutkimusta ja tietoa. [2]

Manyika et al. [5] esittävät hyvin raportissaan kuinka avoin data suhtautuu muihin datatyyppeihin. Siitä käy hyvin ilmi kuinka avoin data on vain osa isompaa datan joukkoa ja miten julkishallinnon datavarannot suhtautuvat avoimeen dataan. Avoimen datan suhteet eri datatyyppien suhteen esitettynä kuvassa 2.2.



**Kuva 2.2.** Avoimen datan suhde muihin datatyypppeihin [mukaillen 5].

EU:n alueella kansallisten datavarantojen avaaminen on alkanut tapahtua 2010-luvun lopulla INSPIRE-direktiivin myötä. INSPIRE-direktiivi 2007/2/EU on Euroopan unionin 15.5.2007 jäsenvaltioille voimaantullut säädös paikkatietoinfrastruktuurin kehittämisestä ja avaamisesta jäsenvaltioiden kansalaisille, yrityksille ja viranomaisille. Säädös velvoitti julkishallintoa direktiivin toimeenpanon mukaisesti avaamaan paikkatietoaineiston katse-lupalvelun ja latauspalvelun vuoden 2011 aikana. [6]

## 2.2 Avoimen datan ominaisuudet

Avoimen datan määritelmä ei ole aivan täysin yksiselitteinen. Datan avoimuuteen vaikuttavat esimerkiksi tekninen saavutettavuus, löydettävyys, maksullisuus ja lisensointiehdot. Aineiston muuttaminen avoimeen suuntaan voidaan kuitenkin toteuttaa pienillä muutoksilla esimerkiksi muuttamalla käyttöehtoja sallivammaksi tai tarjoamalla aineisto kokonaan maksuttomasti. [2] Poikola et al. [2] oppaan mukaan avoimen datan ominaisuudet ovat seuraavat, jotka käydään alaluvuissa läpi.

### 2.2.1 Löydettävyys

Aineiston sekä niiden käyttöehtojen tulisi olla helposti saatavilla ja tunnetussa sijainnissa. Aineiston tulisi olla hakukoneystävällistä, jotta sekä ihmiset että hakurobotit voisivat sen helposti löytää. Aineistossa tulisi olla mahdollisimman vähän fyysisen pääsyn estäviä tekijöitä, kuten esimerkiksi aineiston saatavuutta rajoitettaisiin lomakkeiden täyttämällä



tai jonkun tietyn toimiston käyttämisellä. [7] [2]

Aineiston saaminen myös eri teknologisten taitojen omaaville ihmisille tulisi olla mietittyinä. Käyttöliittymä olisi oltava sellainen, jotta siitä voitaisiin ladata tietokannan tiedot mahdollisimman vaivattomasti. Ohjeistukset tietopyyntöjen tekemiseen API:n kautta sovelluskehittäjiä varten tulisi olla helposti löydettävissä. [7]

### **2.2.2 Kokonaisuus**

Aineisto on saatavissa kokonaisuudessaan, eikä sen käyttöä rajoiteta esimerkiksi vain osaan tietokannasta kerrallaan [2]. Aineistoa ei rajoiteta kuin siinä määrin mikä on lainsäädännön kannalta tarpeellista. Myös metatiedot, jotka selittävät aineiston tietueet on sisällytettävä avoimeen aineistoon. [7]

Mikäli kokonaisuuden määritelmä täyttyy, mahdollistaa se kenen tahansa myös jakaa dataa sekä itseään että muita varten. Kokonaisuuden rajoittamisella voi olla vaikutusta datan kattavaan analysointiin sekä siihen, että API:n kautta tehtävät kyselyt lisääntyvät. [2]

### **2.2.3 Käyttöehtojen tasa-arvoisuus**

Aineiston tulee olla saatavilla kaikille käyttötarkoituksesta riippumatta ja käyttäjällä tulee olla pääsy aineistoon ilman rekisteröintiä. Aineiston käyttöehdot olisi hyvä myös olla toteutettu standardin mukaisilla lisenssiehdoilla. [2]

Käyttöehtojen tasa-arvoisuuteen kuuluu myös, että aineiston käyttö ei ole rajattu vain jollekin tietyille sovelluksille, vaan sen hyödyntäminen on mahdollista sovelluksesta riippumatta. Parhaimmillaan tasa-arvo toteutuu, kun kuka tahansa voi käyttää aineistoa. Olisi myös suotavaa, että aineistoa voisi käyttää ilman tunnistautumista mikäli mahdollista ja aineiston käyttöä ei tarvitsisi perustella. [7]

### **2.2.4 Alkuperäisyys ja ajantasaisuus**

Aineisto on tarjolla mahdollisimman alkuperäisessä muodossaan sekä sellaisella tarkkuudella, että se ei loukkaa tietosuojaa. Arkaluontoinen aineisto on mahdollista tehdä avoimeksi anonymisoimalla, yhdistelemällä ja yleistämällä aineistoja suurta huolellisuutta noudattaen. [2]

Aineiston tulee olla ajantasaista, ja mikäli aineisto on aikaherkkä, tulisi julkaista mahdollisimman nopeasti luontihetkestä. Aineiston reaaliaikainen päivittyminen maksimoi aineistosta saatavan hyödyn, jonka käyttäjä voi siitä saada. [7]

### 2.2.5 Laillinen ja vapaa uudelleenkäytettävyys

Aineisto tulee olla suojattu sellaisilla käyttöehdoilla, jotka mahdollistavat aineiston uudelleenkäytön ja yhdistämisen muihin aineistoihin. Tällaisia käyttöehtoja tarjoavat esimerkiksi Creative Commons ja Open Database –lisenssit. [8] [2]

### 2.2.6 Maksuttomuus

Usein aineistojen julkaisemisesta ja tuottamisesta syntyy kustannuksia, jotka voivat olla suuriakin suhteutettuna aineiston luovuttamisesta perittäviin nimellisiin korvauksiin. Aineiston luovuttamisesta perittävät korvaukset kuitenkin heikentävät sitä kuka on halukas sitä käyttämään. Se voi pahimmillaan, jopa estää aineiston ympärille syntyvää liiketoimintaa. [7]

Parhaimmassa tapauksessa aineisto tulisi olla saatavissa täysin maksutta. Mikäli aineistosta peritään maksua, tulisi tämä olla maksettavissa mahdollisimman helposti esimerkiksi internetissä. Aineisto olisi myös hyvä saada käyttöön mahdollisimman helposti maksun jälkeen esimerkiksi suoraan ladattavana linkkinä tai sähköpostilla. [2]

### 2.2.7 Koneluettavuus

Koneluettavassa muodossa aineisto voidaan lukea koneellisesti, eikä sen tulkintaan tarvita välttämättä ihmislukijaa. Aineisto on jäsennelty myös niin, että se mahdollistaa mahdollisimman helpon koneluettavuuden. [2]

Aineistolla tulisi myös olla myös pysyvä sijainti, jotta se olisi myös saatavilla pitkänkin ajan kuluessa. Mikäli tietoa päivitetään tai poistetaan, tulisi siitä myös ilmoittaa palvelussa. Tiedon päivityksessä tulisi noudattaa version historiointia ja vanhan tiedon arkistointia. [7]

### 2.2.8 Formaatin avoimuus

Aineiston tulisi olla saatavilla sellaisessa yleisessä formaatissa, jonka hallinnointi ja kehitys ei ole yhden yrityksen hallinnassa. Aineiston tulisi olla tarjolla useampia formaatteja ja ohjelmia, joilla aineistoa voidaan hyödyntää tulisi olla useita. Useamman formaatin ja nimenoman maksuttoman formaatin tukeminen mahdollistaa aineiston jakamisen suuremmalle potentiaaliselle käyttäjäjoukolle. [2] [7]

### 2.2.9 Ymmärrettävyys

Aineisto on kuvailtu ja dokumentoitu käyttöesimerkein riittävällä tarkkuudella, jotta aineiston ymmärrettävyys ja uudelleenkäyttäminen helpottuisi. Haittapuolena hyvässä dokumentaatiossa on sen vaatima resurssien tarve, kun dokumentaatiota tehdään. [2]

Avoimen datan ymmärrettävyys korostuu erityisesti julkishallinnon tarjoamassa avoimessa datassa. Tällaisen datan käyttäjinä voi olla hyvin erilaisen teknisen taustan omaavia henkilöitä. Onkin tärkeää, että kaikilla olisi teknisestä taustasta riippumatta helppo pääsy julkishallinnon tarjoamaan avoimeen dataan.

## 2.3 Avoimen datan julkaisu

Avoim data itsessään ei määrittele kuinka tieto on julkaistava. WWW-selaimella on helpos-  
ti saatavilla tietoa suuria määriä tietoa tekemällä hakuja ja selailemalla. Selaimen käyttä-  
minen suurien datamäärien läpikäymiseen on kuitenkin hidasta ja virhealtista. Haastavaa  
onkin valita, miten tieto voidaan ottaa käyttöön eri organisaatioissa. [9]

Avointa dataa voidaan julkaista, niin asiakirjana, kuin helposti koneluettavassa muodos-  
sa. Asiakirjana esitettävään aineistoon käytetään yleensä internet sivustojen tapauk-  
sessa HTML-formaattia ja asiakirjoissa pääsääntöisesti PDF-formaattia. Koneluettavaan  
muotoon käytetään esimerkiksi XML-, CSV-, JSON- ja RDF-formaatteja. [2]

Avoimen datan ylläpitäminen kuuluu sen julkaisijan vastuulle. Eikä esimerkiksi sen päivit-  
tämiseksi taikka julkaisutiheydelle ole asetettu mitään standardeja. Mikäli avointa dataa  
hyödynnetään, olisikin syytä varmistua myös julkaisijan kyvystä ja halusta ylläpitää jul-  
kaistua dataa ajantasaisena.

## 2.4 Avoin data Suomessa

Avoimen datan palveluja on avattu valtion hallinnon toimesta jo vuodesta 2011 lähtien.  
Valtion tietovarantoja on nykyään saatavilla avoimena seuraavista aiheista: maastotieto-  
ja, sää-, ilmasto- ja meridataa ja ilmastomalleja, liikennedatata, ajoneuvotietoja, tilastoai-  
neistoja, julkisia verotietoja, maaperätietoja, talousdataa ja kulttuuriaineistoja. [10]

Vuonna 2013 Valtionvarainministeriö perusti avoimen tiedon ohjelman. Sen tarkoituksen  
oli vauhdittaa ja koordinoita avoimen datan saatavuuden parantamista julkishallinnos-  
sa. Ohjelman tavoitteena on, että kaikki merkittävät julkiset tietovarannot olisi saatettu  
avoimeksi dataksi kaikkien saataville vuoteen 2020 mennessä. [10] Tätä varten avattiin  
vuoden 2014 lopussa Väestorekisterikeskuksen ylläpitämä Avoindata.fi -verkkopalvelu,  
jonka tarkoituksena on kerätä ja saattaa julkisen hallinnon tarjoama avoin data helpos-  
ti saataville yhteen paikkaan [11]. Suurilla kaupungeilla kuten Helsingillä ja Tampereella

on myös omat lähialueisiinsa painottuvat avoimen datan portaalit. Helsingillä on pääkaupungin kattava Helsinki Region Infoshare (<https://hri.fi>) ja Tampereella on Tampereen kaupungin dataportaaali (<https://data.tampere.fi/fi/>).

Suomea pidetään avoimen datan vertailussa yhtenä kärkimaana. Open Knowledge Foundation -järjestön tekemässä avoimen datan indeksivertailussa vuonna 2015 Suomi jakoi viidennen sijan Australian kanssa. Kärkinelikko oli järjestyksessä seuraava: Taiwan, Iso-Britannia, Tanska, Kolumbia. Vertailussa oli mukana 94 maata ja siinä vertailtiin muun muassa kansallisen tilastotiedon, karttojen ja valtion budjetin avoimuutta. [12]

## 2.5 Avoimen datan ongelmat

Vaikka suuntana on avoimempi yhteiskunta, pitää kuitenkin muistaa, että mahdollisia väärinkäytöksiä voi tapahtua. Avoimen tiedon julkaisijan pitää hyvin tiedostaa vastuunsa julkaistavan materiaalin suhteen ja pohtia tarkasti on julkaistava tieto liian yksilöivää. Esimerkiksi Tilastokeskuksen 250 m x 250 m Ruututietokannasta olisi helppo saada yksilöityä tietoa nimenomaan haja-asutusalueilla. Tämän takia Tilastokeskus onkin rajannut haja-asutusalueet pois Ruututietokannasta.

Lähes tarkalleen vuosi sitten uutisoitiin laajalti, kuinka Liikenteen turvallisuusvirasto Trafi oli avannut taksilain uudistuksen yhteydessä ajokorttitietojen hakuun tarkoitetun kuljetajatiedot -palvelun, josta pystyi selvittämään henkilön ajo-oikeuden ja ajokorttiluokan. Palvelusta oli mahdollista selvittää ihmisten tarkkoja syntymäaikoja ja jopa henkilötunnuksia kokonaisuudessaan. Samoin henkilötietojen luovutuskiellossa olevien henkilöiden paikkakuntia oli mahdollista saada selville. Kohun seurauksena Trafín silloinen pääjohtaja päätyi eroamaan. Tämä oli hyvä esimerkki siitä, kuinka hyvää tarkoittavalla palvelulla oli vakavia seurauksia yksityisyyden ja tietoturvan kannalta. [13]

Yhtenä ongelmana voidaan pitää myös datamäärän suuruutta. Koska dataa on tarjolla paljon, vaatii sen hyödyntäminen sopivia työkaluja. Tämän lisäksi käyttäjältä vaaditaan osaamista juuri sen oleellisen datan löytämiseen. Toisaalta suuresta datamäärästä on mahdollista jalostaa tarkempaa yhdistelemällä eri datalähteitä.

## 2.6 Avoimen datan tulevaisuus

Tälle hetkellä ajatellaan, että digimurros on vasta alkuvaiheessa. Joten datapotentiaalin hyödyntäminen on vasta alkutekijöissä. Suomen kuudella suurimmalla kaupungilla; Helsingillä, Vantaalla, Espoolla, Turulla, Tampereella ja Oululla; on tällä hetkellä menossa yhteishanke nimeltään Avoindata ja rajapinnat, jonka tarkoituksena on kehittää yhdessä avoimia ja älykkäitä palveluita [14].

Viime vuosina yleistynyt termi IoT eli Internet of Things, eli suomeksi asioiden internet. Asioiden internet siis käytännössä tarkoittaa erilaisten laitteiden, esimerkiksi pyykinpesu-

koneen kytkeytymistä internettiin, jolloin siitä saadaan tietoa ja sitä voidaan jopa ohjata kännykän kautta. IoT-laitteiden tarjoajilla on vielä hyvin useasti oma pilvipalvelu, johon laitteen täytyy olla yhteydessä, jotta etäohjaus on mahdollista. Samalla käyttäjä antaa luvan valmistajalle kerätä tietoa laitteen käytöstä ja mahdollisesti saa laitteeseensa ohjelmistopäivityksiä.

Avointa dataa hyödyntävistä uusista sovelluksista voidaan mainita erilaiset kotiautomaation laitteet, joilla säädetään kodin lämmitystä. Näitä laitteita hallitaan kännykkäsovelluksen tai verkkokäyttöliittymän kautta ja käyttäjän on mahdollista asettaa haluttu lämpötila tai vaikka luoda lämmitysprofilleja, joka seuraa sisä- ja ulkolämpötiloja.

## **2.7 Esimerkki avoimen datan saatavuudesta ja sen esittämisestä**

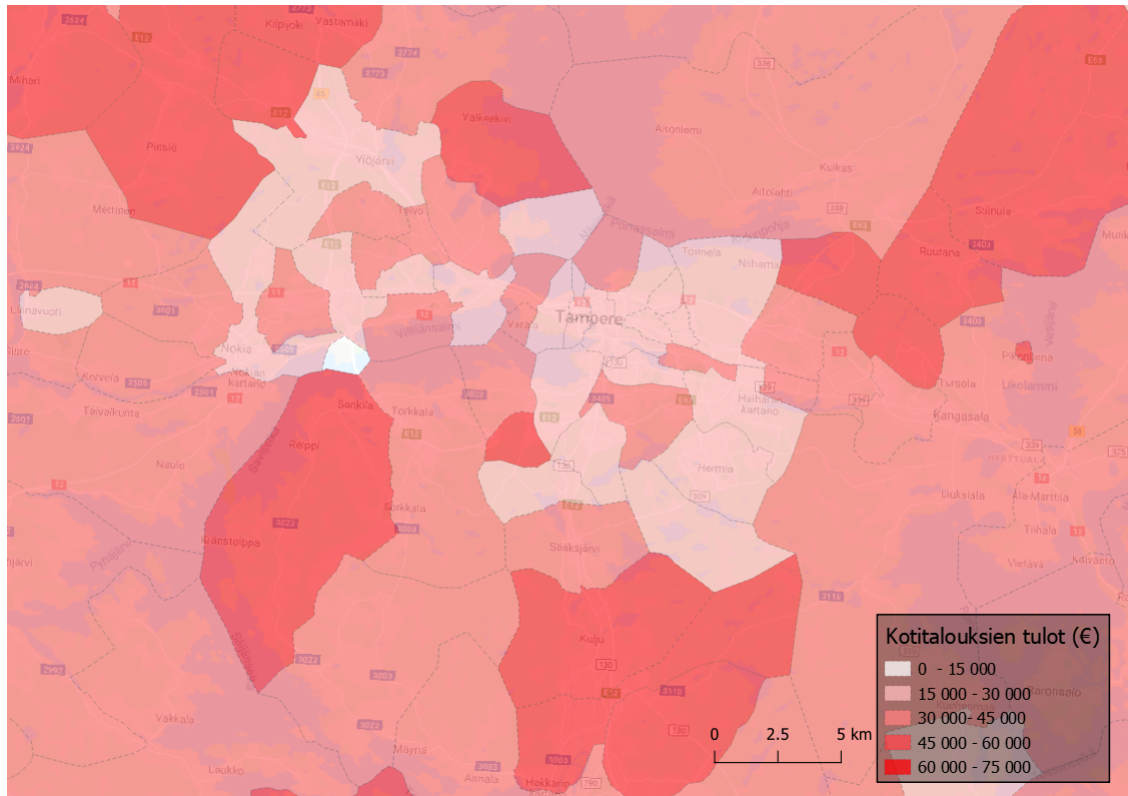
Esimerkkiä varten koitin etsiä mahdollisimman tarkkaa ja ajantasaista aineisto. Tilastokeskukselta löysin ruututietokannan, jossa aineistoa oli tarjolla 250m x 250m aina 5 km x 5 km kokoisina karttaruudukoina. Harmikseni sain huomata, että tämä ruututietokanta oli tarjolla vain maksua vastaan. Tilastokeskuksella oli myös tarjolla PAAVO-palvelu, jossa sama tieto on tarjolla kuin ruututietokannassa mutta PAAVO-palvelussa se on tarjolla vain postinumeroalueittain. Tämän esimerkin kannalta postinumeroalueet ovat riittävä tarkkuus, sillä tässä esimerkissä on tarkoitus näyttää, kuinka avointa dataa on konkreettisesti saatavilla ja kuinka sitä on mahdollisuus hyödyntää.

PAAVO-palvelussa on tarjolla paljon tietoa yksittäisestä alueesta tallennettuna eri muuttujiin. Eri muuttujia tiedolle oli yli 100 kappaletta ja ne sisältävät tietoa kotitalouksien koosta, ikäjakaumasta, koulutuksesta, taloudesta ja työllisyydestä. Tilastokeskuksella oli palvelussaan esillä myös ohjeet ja käyttöehdot datan käyttämiseen, niin kuin luvussa 2.2 avoimen datan hyviltä ominaisuuksilta vaadittiin. PAAVO-palvelu on saatavilla osoitteessa: <https://www.stat.fi/tup/paavo/index.html>.

PAAVO-palvelussa avoin tieto on saatavilla PxWeb-tietokantana tai sitten WFS- ja WMS-rajapintojen kautta. WFS- ja WMS-rajapinnan hyödyntäminen vaatii jonkun näitä tukevan sovelluksen. Tässä esimerkissä on hyödynnetty WMS-rajapintaa QGIS-ohjelmistolla. QGIS-ohjelmisto on vapaasti saatavilla lähes kaikille käyttöjärjestelmille paikkatiedon esittämiseen ja hallintaan tarkoitettu ohjelmisto. Kuvassa 2.3 on esitetty valmis kuvaus QGIS-ohjelmistolla toteutettuna Tampereen kotitalouksien mediaanituloista postinumeroalueittain vuonna 2016.

Aluksi QGIS-ohjelmistolla on haettu Google Maps -palvelusta kartta taustakartaksi ja muutettu kordinaatisto vastaamaan Suomessa käytössä olevaa kartan tasokordinaattijärjestelmää, ETRS-TM35FIN. Tämän jälkeen Tilastokeskuksen PAAVO-palvelusta WMS-rajapinnan kautta on haettu tiedot QGIS-ohjelman omilla työkaluilla postinumeroalueittain uudelle tasolle. QGIS-ohjelmistolla eri tiedot voidaan lisätä eri tasoille, jolloin tiedot pysy-

vät järjestyksessä ja näiden muokkaus on helpompaa. PAAVO-palvelussa eri tiedot on jaettu eri muuttujiin. Tilastokeskuksella on tarjolla käyttöopas ja muuttujaluettelo PAAVO-palvelulle. Muuttujaluettelosta täytyi tarkastaa oikea muuttujan nimi halutulle tiedolle, tässä tapauksessa kotitalouksien mediaanituloille. Ohjelmistossa valittiin muuttujalle sopivat arvoalueet ja tämän perusteella postinumeroalueet värjättiin arvon mukaan.



**Kuva 2.3.** Tampereen alueen kotitalouksien mediaanitulot postinumeroalueittain vuonna 2016 [15]. Vaaleamman punaiset alueet ovat alemman tulotason alueita ja punaisemmat alueet korkeamman tulotason alueita.

## 3 SÄHKÖVERKKO

Suomessa ensimmäiset sähköverkot näkivät päivänvalon 1800-luvun lopulla. Tampereella Finlaysonin tehdas otti ensimmäisenä sähköllä toimivan valaistuksen käyttöön. Aluksi sähköverkot ja sähköntuotanto oli sidoksissa teollisuuslaitoksiin. 1900-luvun puolivälissä, sähkölaitteiden yleistyessä, nykyisen kaltainen sähköverkko on saanut alkunsa.

Sähköverkon pääasiallinen tarkoitus on yhdistää sähköntuotantopaikat ja sähkönkulutuspaikat toisiinsa. Suomessa sähköverkko toimii vaihtojännitteellä, jonka taajuus on 50 Hz. Sähköverkossa tuotannon ja kulutuksen on aina pysyttävä tasapainossa toisiinsa nähden. Mikäli tuotanto on suurempaa kuin kulutus lähtee taajuus nousemaan ja vastaavasti kulutuksen ollessa suurempaa lähtee taajuus laskemaan.

Suomen sähköverkko voidaan karkeasti jakaa kahteen osaan: kantaverkkoon ja jakeluverkkoon. Näistä vielä tarkemmin jäljempänä omissa alaluvuissaan. Sähköverkko on myös yhteydessä muihin pohjoismaihin, pois lukien Islanti, sekä Viroon ja Venäjään. Sähköverkon yhteys muiden maiden sähköverkkoon mahdollistaa sähkön tuonnin ja viennin. Suomessa sähköenergiatase on alijäämäinen, eli tänne täytyy siirtää sähköenergiaa muista verkkoon yhteydessä olevista maista. Tuontisähkön osuus Suomen sähkönkulutuksesta vuonna 2018 oli noin 30 % [16].

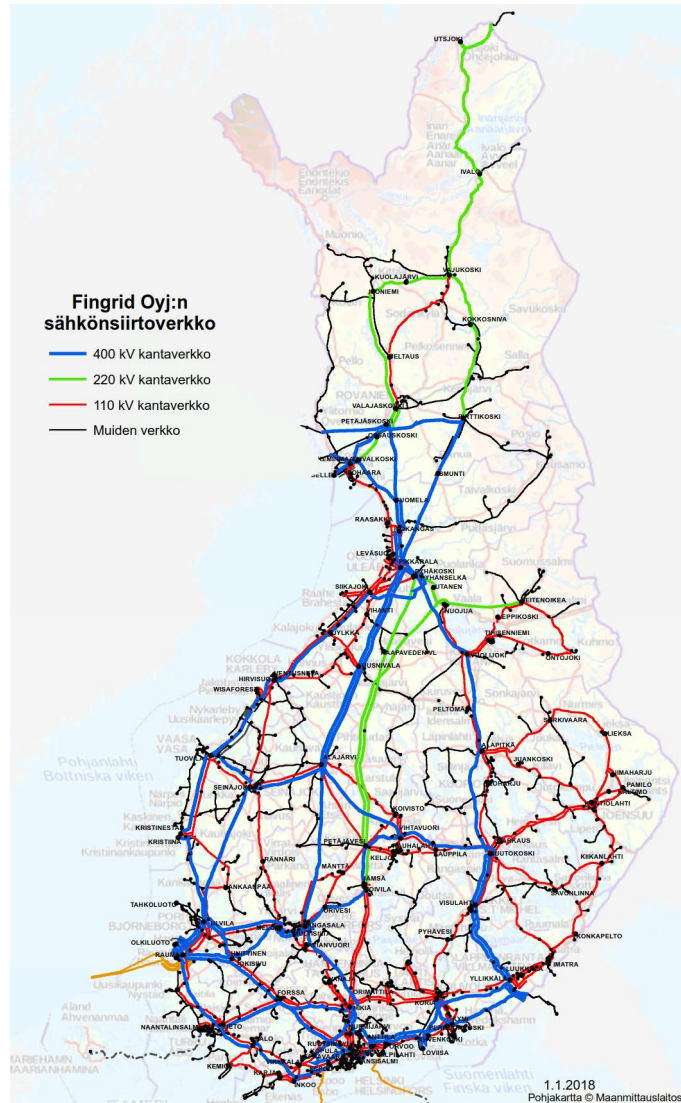
Sähköenergian kulutus Suomessa vuonna 2018 oli noin 86 TWh. Suurin osa tästä kulutuksesta koostuu teollisuuden sähkönkäytöstä ja kiinteistöjen sähkölämmityksestä. Sähköverkon kulutuksen keskiteho oli vuonna 2018 keskimäärin 9792 MW ja huipputeho 14 062 MW. [16]

### 3.1 Kantaverkko

Kantaverkko on Fingrid Oyj:n hallinnassa ja sen omistavat suurimmaksi osaksi Suomen valtio ja suuret vakutusyhtiöt [17]. Fingrid Oyj on saanut alkunsa vuonna 1997 kun liiketoimintakaupalla kantaverkon kehitys ja hallinnointi haluttiin keskittää yhteen yhtiöön. Fingrid Oyj:llä on lain määräämä vastuu koko Suomen sähköjärjestelmästä ja se osallistuu eurooppalaisen ENTSO-E:n toimintaan, jonka tarkoituksena on kehittää yhteiseurooppalaisia energiemarkkinoita [18].

Kantaverkko on maamme sähkönsiirron selkäranka. Kantaverkko kattaa jännitetasot 110 kV:sta aina 400 kV:iin ja se on suunniteltu silmukkamallisesti, jonka tarkoituksena on pa-

rantaa sähkön toimitusvarmuutta siirtopaikkoihin. Kuvassa 3.1 on esitettyä vuoden 2019 kantaverkko. Siitä huomataan kantaverkon silmukkamallinen suunnittelu, joka kattaa lähes koko verkon aivan pohjoista verkkoa lukuun ottamatta. Kantaverkko käsittää tällä hetkellä yli 14 000 km siirtolinjaa sekä yli 100 sähköasemaa [18].



**Kuva 3.1.** Fingrid Oyj:n kantaverkko vuonna 2019.

## 3.2 Jakeluverkko

Jakeluverkot ovat sähköverkkoyhtiöiden vastuulla ja se kattaa jännitetasot 0,4 kV – 110 kV Jakeluverkkoon liittyvät myös muut sähköverkon osat kuten sähköasemat, jakelu-muuntamot ja pienjänniteverkko. Edellä mainittujen komponenttien lisäksi jakeluverkkoon sisältyy myös käyttöä tukevia laitteistoja ja toimintoja kuten verkon suojalaitteet, tiedon-siirtolaitteet, käytönvalvontajärjestelmät ja tietojärjestelmät [19].

Toisin kuin kantaverkossa, käytetään jakeluverkossa säteittäistä verkon muotoa. Kaupun-geissa verkko voi muodoltaan olla silmukkamainen mutta käyttö tapahtuu kuitenkin sä-



teittäisesti. Säteiltein käytettävään jakeluverkkoon on suurimmaksi osaksi päädytty kustannussyistä. Silmukkamaisesti käytettynä sähköverkko tarvitsee huomattavasti kalliimman suojauksen kuin säteiläinen verkko. Lisäksi syrjäalueilla käyttökohteiden etäisyydet lähes pakottavat suunnittelemaan verkko säteiläiseksi. [20]

## 4 SÄHKÖVERKKOTOIMINTA

Sähköverkkotoiminta Suomessa on tarkasti säänneltyä ja luvanvaraista toimintaa. Toimilupien myöntäminen on Energiaviraston vastuulla, joka myös valvoo sähköverkkotoimintaa. Sähkömarkkinalaki velvoittaa sähköverkkoyhtiöt toimittamaan hyvälaatuista sähköä ylläpitämällä ja kehittämällä sähköjakeluverkkoa asiakkaiden tarpeiden mukaan. Sähkömarkkinalaki ottaa myös kantaa sähkön hinnoitteluun. Sähköverkkoyhtiöt ovat oikeutettuja kohtuulliseen tuottoon, jossa otetaan huomioon riittävään toimintavarmuuteen ja tehokkuuteen tähtäävät kannustimet. [21]

Sähköverkkotoiminnassa vallitsee luontainen monopoli, joten siinä ei ole luontaista kilpailua, joka vaikuttaisi hinnoitteluun ja laadun kehittämiseen. Asiakkaat tietysti odottavat halpaa hintaa ja hyvälaatuista sähköä. Sähkön hinta koostuu kolmesta tekijästä: sähköenergiasta, sähkönsiirrosta ja veroista, jakautuen karkeasti kolmasosiin kaikkien kesken. [21]

Euroopan parlamentin ja neuvoston sisämarkkinadirektiivi (2003/54/EY) aiheutti sen, että vuonna 2005 Suomessa päätettiin uudistaa sääntelyjärjestelmä. Sääntelyjaksot muuttuivat nyt neljän vuoden mittaisiksi. Sähköverkkoyhtiöille lisättiin kannustimia laatuun, tehostamiseen, investointiin ja toimitusvarmuuteen. Jokaisen sääntelyjakson lopussa myös mietitään, miten sääntelyä voitaisiin edelleen kehittää ja pitääkö sääntelyä muuttaa seuraavalle sääntelyjaksolle. [21]

### 4.1 Toimitusvarmuusvaatimukset

Toimitusvarmuusvaatimukset saivat alkunsa vuosituhannen vaihteen suurista myrskyistä sekä vuosikymmen myöhemmin vallinneista Astasta (2010) ja Tapanista (2011). Pisimmät sähkökatkot Astan myrskyssä olivat yli kuukauden mittaisia ja Tapanin myrskyssä lähes puoli miljoonaa sähkönkäyttäjää oli ilman sähköä tunti- ja päiväkausia. [22]

Toimitusvarmuusvaatimuksiin on kirjattu tällä hetkellä käytössä olevat maksimikeskeytysajat taajamassa 6 tuntia ja haja-asutusalueella 36 tuntia. Toimitusvarmuusvaatimukset ovat kirjattu sähkömarkkinalakiin ja täten sähköverkkoyhtiöillä velvoite täyttää toimitusvarmuusvaatimukset. Vuonna 2003 tuli käyttöön myös laissa säädetty vakiokorvaukset toimituksen keskeytyksistä. [22]

## 4.2 Sähköverkon kehittäminen

Sähköverkossa verkon osien pitoajat ovat yleensä hyvin pitkät, 30 - 50 vuotta [19], joten tämän hetken kehittämispäätöksillä voi olla suuria vaikutuksia taloudellisuuden suhteen pitkällä aikavälillä. Parasta tietysti olisi, että tulevaisuuden sähkönkäyttö pystyttäisiin mahdollisimman tarkasti mallintamaan tai ennustamaan.

Vuonna 2007 EU vahvisti 20/20/20 -tavoitteet. Nämä luvut tarkoittavat, että vuonna 2020 20 % Euroopan energian kulutuksesta olisi katettava uusiutuvilla energianlähteillä, kasvihuonekaasupäästöjen määrän tulee pienentyä 20 % vuoden 1990 tasosta ja energiatehokkuuden on parannuttava 20 % vuoden 2005 energiatehokkuuteen nähden [23]. Tämä yhdessä kiristyneiden sähkön toimitusvarmuusvaatimusten kanssa on asettaa sähköverkkoyhtiöille haasteita verkon kehitykseen.

Hajautetun tuotannon lisääntyminen ja verkkoon kytkeminen asettaa myös haasteita sähköverkon kehittämiseen. Sähköverkossa sallittu jännitteen vaihteluväli on  $\pm 10$  %. Tällöin, jos 400 V:n pienjänniteverkon haarassa on pientuotantoa, voi pientuotannon liityntäpaikalla jännite nousta yli sallitun. Ongelman voisi ajatella korjaantuvan laskemalla vain jakelumuuntamon jännitettä. Jakelumuuntamon jännitteen laskulla voi kuitenkin olla vaikutusta verkon haaraan, jossa ei ole pientuotantoa. Tällaisessa verkon haarassa jännite voi taas laskea alle sallitun. Tämän takia säteittäisesti toimiva jakeluverkko toimii käytännössä oikein vain yhteen suuntaan [20].

## 4.3 Smart grids

Smart grids eli älykkäät sähköverkot yhdistävät automaatio- ja tietoliikenneteknologiaa osaksi sähköverkkoa. Smart gridin myötä älykkäät sähkömittarit ovat yleistyneet ja ovat nykyään pakollisia. Nykyajan älykkäillä sähkömittareilla on mahdollista seurata kulutuspaikan tietoja lähes reaaliaikaisesti mutta käytännössä mittareiden luku tapahtuu kerran päivässä.

Tulevaisuuden sähköverkkojen ajatellaan koostuvan pelkästään älykkäistä sähköverkoista. Tulevaisuuden sähköverkkojen ajatellaan koostuvan paljon hajautetusta ja vaihtelevasta tuotannosta sekä uusista sähköverkkoon liitetyistä komponenteista kuten sähköautoista ja virtuaalisista voimalaitoksista. Jotta älykäs sähköverkko toimii luotettavasti, tarvitaan nopeita tietoliikenneyhteyksiä. 5G-verkkojen kehityksessä älykkäiden sähköverkkojen tarpeilla on ollut myös suuri rooli. [24]

## 4.4 Sähköverkkotoiminnassa käytettävät tietojärjestelmät

Sähköverkkoyhtiöillä on käytössä tukenaan useampia tietojärjestelmiä, kuten käytöntu-  
kijärjestelmä, verkkotietojärjestelmä ja asiakastietojärjestelmä. Nämä kaikki järjestelmät

ovat hyvin vahvasti kytköksissä toisiinsa ja käyttävätkin toiminnan kannalta oleellisia yhteisiä tietokantoja. Tietokantojen etuna on se, että sitä kautta saadaan helposti päivitettyä tieto moneen eri käytössä olevaan järjestelmään kerralla.

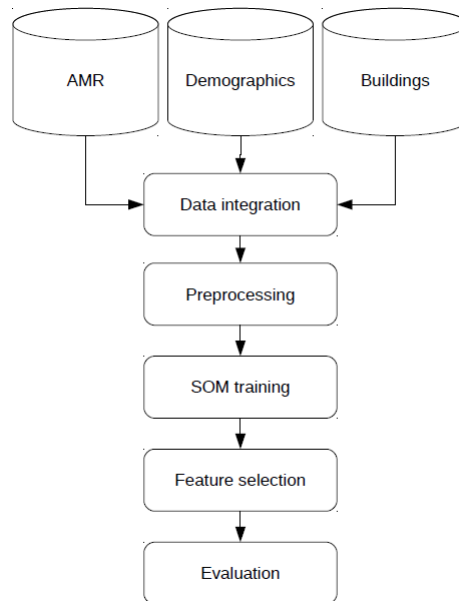
Käytöntukijärjestelmällä seurataan verkon tilaa ja voidaan hallita verkon aktiivisia komponentteja. Sen avulla saadaan tieto myös vikatilanteesta ja useasti järjestelmä osaa myös hoitaa tarvittavat toimenpiteet vikatilanteessa. Verkkotietojärjestelmä on tarkoitettu omaisuuden hallintaan ja verkostosuunnitteluun. Järjestelmästä saadaan tietoa käytöstä ja sijainnista yksittäisen komponentin tarkkuudella. Asiakastietojärjestelmä sisältää tiedot asiakkuuksista ja kulutuspaikoista. Tämän avulla voidaan seurata yksittäisen sähkönkäyttäjän sähkönkulutustietoja.

## 5 AVOIMEN DATAN HYÖDYNTÄMINEN SÄHKÖVERKKOTOIMINNASSA

Sähkön käytön ennustaminen ja verkoston suunnittelu on hyvin pitkälti perustunut tällä hetkellä kulutusprofiileihin [25]. Kulutusprofiilit ovat erilaisille sähkökäyttäjille tehtyjä malleja, kuten teollisuudelle ja kotitalouksille, joiden avulla on pyritty laskemaan ja mitoittamaan sähköverkkoa mahdollisimman tarkasti. Kulutusprofiilien etuna on, että eri kuluttajan parametrejä vaihtamalla voidaan helposti ja nopeasti simuloida erilaisia sähköverkon kulutuskenaarioita. Nykyään kuitenkin hajautetun pientuotannon lisääntyessä, nämä vanhat kulutusprofiilit eivät enää vastaa todellista sähkökuluttajan kulutusta. Samoin etäluettavien mittarien ansiosta saadaan tarkempaa tietoa eri kulutuksista, joiden avulla on mahdollista luoda tarkempia kulutusmalleja [25]. Tämän vuoksi onkin mahdollista jo nyt, että näitä kulutusprofiileja päivitetään vastamaan tämän hetkistä kuluttajaa tai sovelletaan vaihtoehtoisia tapoja ennustaa sähkökulutusta.

Itä-Suomen yliopistossa on tehty tutkimus [25], jossa yhdistellään etäluettavien mittareiden tietoa ja avointa dataa. Tietojen yhdistämisen tarkoituksena oli luoda uudenlaisia kulutusprofiileja vastamaan tämän hetken sähkökäyttäjiä. Etäluettavasta kulutusmittarista saatiin käyttöpaikan kulutustiedot ja sen hetkinen säätötila. Avoin data oli Tilastokeskuksen ruututietokannan 250 m x 250 m ruututietoaineistoa, josta hyödynnettiin sosiodemografista aineistoa. Tutkimuksessa käytettiin myös avoimesti saatavilla olevia rakennustietoja, joista saatiin selville rakennuksen lämmitysmuoto. Tutkimuksessa rajattiin ulkopuolelle teollisuuden ja kaupanalan käyttöpaikat ja keskityttiin vain asuinkiinteistöjen sähkökäyttöön.

Kuvassa 5.1 on esitetty tutkimuksen datan keruun ja suodatuksen prosessikaavio. Aluksi yhdistettiin käyttöpaikkojen tiedot, sosiodemograafiset ruututiedot ja rakennusten tiedot. Yhdistämisen pohjana käytettiin sosiodemografista ruututietoa. Tämä jälkeen mittareiden kulutustietoihin sovellettiin lämpötilan korjauskertoimia. Lopuksi SOM- ja KNN-algoritmien avulla saatiin ruututietokannan ruudulle erilaisia mahdollisia kulutusmalleja. SOM eli itseorganisoiuva kartta on algoritmi, jolla voidaan esittää sekä ryhmitellä monimutkaisia ja epälineaarisia riippuvuussuhteita sisältäviä alkioita, kuitenkin likimäärin säilyttäen alkioiden alkuperäiset ominaisuudet. KNN-algoritmilla taas luodaan  $k$ :n lähimmistä arvoista yksi estimaattiarvo.



**Kuva 5.1.** Itä-Suomen yliopiston tutkimuksen prosessikaavio. [25]

Tutkimuksessa todettiin, että mallinnettu kulutus vastasi todellista kulutusta melko hyvin. Sen vahvuutena oli kuitenkin kulutusprofiilien luonti automatisoidusti sosiodemografisten tietojen ja rakennustietojen pohjalta. Tutkimuksessa ehdotettu malli tuotti useamman kulutusprofiilin eri lähtötiedoilla, joka mahdollistaisi sähköverkkoyhtiön valintaa päätöksenteossa tarkkuuden ja lähtötietojen määrällä. Kulutusprofiilien ruudukointi pieniksi alueelliseksi tiedoiksi koettiin myös merkittäväksi sähköverkkoyhtiöiden hyödyntämiseen.

Julkishallinnon tarjoamasta avoimesta datasta noin 80% sisältää jonkinlaisen viittauksen paikkatietoon ja se on usein esitetty pisteenä tasokordinaatistossa [26]. Tilastokeskuksellakin on saatavilla hyvin tarkkaa ruudukoitua paikkatietoa niin kuin luvussa 2.7 esitettiin. Sähköverkkoyhtiöiden käytössä olevat verkkotietojärjestelmät sisältävät myös suuren määrän yhtiön omaa aineistoa, joka hyvin pitkälti on myös sidottu paikkatietoon. Joten paikkatietoa sisältävä avoin data olisi helpoiten sähköverkkoyhtiöiden hyödynnettävissä.

Toisessa Itä-Suomen yliopistossa tehdyssä tutkimuksessa [27] tutkittiin, kuinka sosioekonomiset tekijät ovat vaikuttaneet sähköhybridiautojen hankkimiseen. Tutkimuksessa hyödynnettiin myös Tilastokeskuksen ruututietokannan aineistoa. Tutkimus noudatteli datan keruun ja suodattamiseen suhteen hyvin pitkälti samanlaista prosessia kuin aiemmin esitellyssä tutkimuksessa. Tässä tutkimuksessa karttaruudukoiden kotitaloudet oli jaettu viiteen eri ryhmään sosioekonomisten piirteiden mukaan.

Tutkimuksessa todettiin, että korkeamman tulotason kotitaloudet ovat herkempiä hankki-  
maan sähköhybridiauton. Toisaalta kotitalouksien sijainnilla on myös vaikutusta ja onko kotitaloudessa useampia autoja kuin yksi. Tuloksista olisi myös saatu merkityksellisempiä, mikäli olisi otettu huomioon myös ruudukon alueen muut autot ja suhteutettu sähköhybridiautojen määrä niihin. Ryhmien määrän olisi tullut olla myös isompi, jolloin tuloksiin olisi tullut enemmän vaihtelua alueittain. [27]

Kuvassa 5.2 on esitettyä tutkimuksen karttaruutuihin jaettu valmis aineisto. Kuvasta selviää hyvin, kuinka se on käytettävissä karttapohjan päällä ja olisi valmiiksi hyödynnettävissä kun ennustetaan kulutuksen lisääntymistä jollakin tietyllä alueella. Tämän lisäksi tulevaisuudessa älykkäiden sähköverkkojen ja sähköautojen akustojen kehittyessä, voidaan sähköautot mieltää myös energiavarastoina, joista voidaan vapauttaa sähköenergiaa verkkoon kulutushuippuja tasaamaan [27]. Tutkimuksen tuloksia voidaan myös hyödyntää, kun mietitään uusien latausasemien rakentamista.



**Kuva 5.2.** Kotitalouksien hankkimat sähköhybridiautot ryhmittäin. Tummempi väri kuvaa suurempaa sähköhybridien määrää per kotitalous. [27]

## 6 YHTEENVETO

Tässä työssä oli tarkoituksena tarkastella, minkälaista avointa dataa on saatavilla ja kuinka tätä dataa olisi mahdollista hyödyntää sähköverkkotoiminnassa. Työssä tutustuttiin avoimen datan ominaisuuksiin ja käytiin läpi sähköverkkoa ja sähköverkkotoimintaa.

Avoimen datan hyödyntäminen sellaisenaan ei ole tarkoituksellista. Sen järkevä ja kustannustehokas hyödyntäminen vaatii käyttäjältä myös jonkun verran teknistä osaamista. Avoimesta datasta hyvin suuri osa sisältää myös paikkatietoa. Toiminnassa, jossa paikkatiedolla on merkitystä, on avoimella datalla arvoa. Avoimen datan esimerkissä nähtiin konkreettisesti, kuinka avointa dataa oli paikkatiedon avulla mahdollista hyödyntää.

Sähköverkkojen kehitys on pitkälti nojannut vanhoihin malleihin. Sähköverkkotoiminnan kannalta olisi järkevää saada ennustettua kuinka sähköenergian kulutus tulee muuttumaan tulevaisuudessa sillä sähköverkkojen komponenttien pitoajat ovat pitkät. Hyvä kulutuksen ennustettavuus mahdollistaa sähköverkkoyhtiöille paremman taloudellisen toimintaympäristön. Sähköverkkoyhtiöillä käytössä olevat ohjelmistot mahdollistavat kulutuspaikkojen ja komponenttien sijainnin tietämisen tarkasti.

Sähköverkkotoiminnan kannalta avoimesta datasta saataisiin hyötyä juuri kulutuksen ennustukseen. Avoimen datan avulla ja nykyisillä etäluettavilla mittareilla on mahdollista luoda tarkempia kulutusmalleja, jotka perustuvat sosiodemografisten muuttujien ja energiankulutuksen yhteneväisyyksiin. Avoimen datan hyödyntäminen tällaisessa tarkoituksessa vaatii myös sen hyödyntäjältä datan käsittelytaitoja. Avointa dataa on mahdollista hyödyntää myös esimerkiksi sähköautojen tapauksessa latausinfrastruktuurin kehittämiseen.

Tulevaisuudessa täyssähköautojen lisääntyessä näiden alueellisten kappalemäärien selvittäminen avoimen datan avulla ja niiden hyödynnettävyys energiavarastona älykkäissä sähköverkoissa vois olla yksi jatkotutkimuksen kohde. Toisaalta tällä hetkellä täyssähköautojen hankintaan voi sosioekonomisten muuttujien lisäksi vaikuttaa myös ideologiset syyt.



## LÄHTEET

- [1] I. Salo. *Big data & pilvipalvelut*. Jyväskylä: Docendo, 2014.
- [2] A. Poikola, P. Kola ja K. A. Hintikka. *Julkinen data - johdatus tietovarantojen avaamiseen*. 2010. URL: <http://urn.fi/URN:ISBN:978-952-243-146-2> (viitattu 17.02.2019).
- [3] Gartner warns of big data security problems. *Network Security* 2014.6 (2014), 20. ISSN: 1353-4858. DOI: [https://doi.org/10.1016/S1353-4858\(14\)70062-5](https://doi.org/10.1016/S1353-4858(14)70062-5). URL: <http://www.sciencedirect.com/science/article/pii/S1353485814700625>.
- [4] S. Chignard. *A brief history of Open Data*. 2013. URL: <http://parisinnovationreview.com/articles-en/a-brief-history-of-open-data> (viitattu 14.03.2019).
- [5] J. Manyika, M. Chui, P. Groves, D. Farrell, S. Van Kuiken ja E. A. Doshi. *Open data: Unlocking innovation and performance with liquid information*. 2013.
- [6] EUROOPAN PARLAMENTIN JA NEUVOSTON DIREKTIIVI 2007/2/EY, annettu 14 päivänä maaliskuuta 2007, Euroopan yhteisön paikkatietoinfrastruktuurin (INSPIRE) perustamisesta. *Euroopan unionin virallinen lehti* L 108/1 (2007). URL: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32007L0002> (viitattu 05.03.2019).
- [7] L. Lessig. *Ten Principles For Opening Up Government Information*. 2017. URL: <https://sunlightfoundation.com/policy/documents/ten-open-data-principles/> (viitattu 21.03.2019).
- [8] D. Dietrich, J. Gray, T. McNamara, A. Poikola, R. Pollock, J. Tait ja T. Zijlstra. *The Open data handbook*. URL: <http://opendatahandbook.org/guide/en/> (viitattu 19.03.2019).
- [9] E. Hyvönen. *Semanttinen web: Linkitetyn avoimen datan käsikirja*. Tallinna: Gaudemus, 2018.
- [10] A. Kauhanen-Simanainen ja M. Suurhasko. *Avoimesta datasta innovatiiviseen tiedon hyödyntämiseen: Avoimen tiedon ohjelman 2013–2015 loppuraportti*. Helsinki, 2015. URL: <https://vm.fi/julkaisu?pubid=6902> (viitattu 21.02.2019).

- [11] Väestörekisterikeskus. *Mitä on avoin data?* URL: <https://www.avoindata.fi/fi/opas/mita-on-avoin-data> (viitattu 21. 02. 2019).
- [12] *Global Open Data Index*. URL: <https://index.okfn.org/place/> (viitattu 05. 12. 2019).
- [13] M. Julku. Kommentti: Trafi julkaisi suomalaisten ajokorttitiedot netissä - Tämän kaiken ehdin saada kavereistani selville. *Iltaalehti* (10. joulukuuta 2018). URL: <https://www.iltalehti.fi/autouutiset/a/68a9e5de-8e06-4986-aa87-d070d91b761d> (viitattu 05. 12. 2019).
- [14] M. Siukonen. *Dataliiketoiminnan tulevaisuus*. URL: <https://www.databusiness.fi/fi/blogi/dataliiketoiminnan-tulevaisuus-takanamme-vasta-muutoksen-murto-osa/> (viitattu 12. 11. 2019).
- [15] Tilastokeskus. *Postinumeroalueittainen avoin tieto*. (Viitattu 16. 11. 2019).
- [16] *Sähkön kulutus ja tuotanto*. URL: <https://www.fingrid.fi/sahkomarkkinat/kulutus-ja-tuotanto/> (viitattu 30. 10. 2019).
- [17] *Osakkeet ja osakkeenomistajat*. URL: <https://www.fingrid.fi/sivut/sijoittajat/osakkeet-ja-osakkeenomistajat/> (viitattu 12. 11. 2019).
- [18] *Fingrid -esittely*. URL: <https://www.fingrid.fi/sivut/yhtio/esittely/> (viitattu 12. 11. 2019).
- [19] E. Lakervi ja J. Partanen. *Sähkönjakelutekniikka*. Helsinki: Otatieto, 2008.
- [20] J. Elovaara ja L. Haarla. *Sähköverkot I*. Helsinki: Otatieto, 2011.
- [21] J. Partanen, S. Viljainen, J. Lassila, S. Honkapuro, K. Salovaara, S. Annala ja M. Makkonen. *Sähkömarkkinat - opetusmoniste*. Lappeenranta, 2017.
- [22] J. Partanen. *Sähkönsiirtohinnot ja toimitusvarmuus*. 2018. URL: <http://urn.fi/URN:ISBN:978-952-327-356-6> (viitattu 21. 10. 2019).
- [23] *Overall progress towards the European Union's '20-20-20' climate and energy targets*. URL: <https://www.eea.europa.eu/themes/climate/trends-and-projections-in-europe/trends-and-projections-in-europe-2016/1-overall-progress-towards-the> (viitattu 19. 12. 2019).

- [24] C. Wietfeld, A. A. Cardenas, H. Chen, P. Popovski ja V. W. S. Wong. Smart Grids. *IEEE Wireless Communications* 24.2 (huhtikuu 2017), 8–9. ISSN: 1558-0687. DOI: 10.1109/MWC.2017.7909091.
- [25] J. Saarenpää, M. Kolehmainen, M. Mononen ja H. Niska. A data mining approach for producing small area statistics-based load profiles for distribution network planning. Teoksessa: *2015 IEEE International Conference on Industrial Technology (ICIT)*. Maaliskuu 2015, 1236–1240. DOI: 10.1109/ICIT.2015.7125266.
- [26] R. Sieber. Open data and geospatial. Teoksessa: *The state of open data: Histories and horizons*. Cape Town ja Ottawa: African Minds ja International Development Research Centre, 2019, 137–150.
- [27] J. Saarenpää, M. Kolehmainen ja H. Niska. Geodemographic analysis and estimation of early plug-in hybrid electric vehicle adoption. *Applied Energy* 107 (2013), 456–464. ISSN: 0306-2619. URL: <http://www.sciencedirect.com/science/article/pii/S0306261913001840> (viitattu 18.02.2019).